Risk-Averse Access Point Selection in Wireless Communication Networks

Wann-Jiun Ma, *Member, IEEE*, Chanwook Oh, Yang Liu, Darinka Dentcheva, and Michael M. Zavlanos, *Member, IEEE*

***Abstract*—This paper considers the problem of selecting the optimal set of access points and routing decisions in wireless communication networks. We consider networks that are subject to uncertainty in the wireless channel, e.g., due to multipath fading effects, and formulate the problem as a risk-averse network flow problem with binary variables corresponding to the status of the sinks, namely, selected or not. Risk measures capture low-probability but high-cost events and, when used for stochastic optimization, they produce solutions that are more reliable compared to mean-value formulations and less conservative than worst-case approaches. By relaxing the integer constraints, we reformulate the problem as a linear optimization problem, which we solve in a distributed way using the ADAL (Accelerated Distributed Augmented Lagrangian) method that was recently developed by the authors to solve optimization problems with convex separable objectives and linear coupling constraints. We present numerical simulations and experimental results using low-power wireless radios that demonstrate the ability of the proposed method to effectively deal with large variations in the quality of the wireless channel.**

***Index Terms*—Optimal wireless networking, distributed optimization, risk-averse optimization.**

I. Introduction

WIRELESS networks are used in many applications involving environmental monitoring, infrastructure inspection, and health-care monitoring, to name a few. The goal in these applications is to enable reliable information delivery from source nodes to a user via a set of access points that can be stationary nodes or mobile robots acting as information sinks. The focus of this paper is to develop new distributed methods to select the optimal access point locations and routing decisions in wireless sensor networks subject to uncertainty and risk in the wireless channel.

Early communication models have typically relied on constructs from graph theory, with proximity graphs gaining the most popularity. This is consistent with early approaches to wireless networking that used disk models to abstract the physical layer [1]. In this context, communication becomes equivalent to topological connectivity, defined as the property of a graph to transmit information between all pairs of its nodes. Control of topological connectivity has typically relied on controlling the Fiedler value of the underlying graph. This approach has been widely used especially in the control of robotic networks due to its simplicity [2]–[8]. Nevertheless, in practice, the above graph-based communication models turn out to be rather conservative, since proximity does not necessarily imply tangible and reliable communication [9]. A simple, yet effective, modification is to use graph models that associate weights to links used to capture either the signal strength [10], or the packet error probability of the link [11]. When using reliabilities as link metrics it is possible to model routing and scheduling as optimization problems that accept link reliabilities as inputs [12], [13]. Relevant optimization methods to analyze and design wireless communication networks can also be found in [14]–[19].

A key challenge akin to optimal wireless networking problems is the presence of uncertainty in wireless transmissions. In fact, due to shadowing and fading effects, channel strength measurements can vary significantly for stationary networks and even more for mobile ones [20]–[22]. This uncertainty necessitates the development of suitable optimization problems and effective stochastic optimization methods to ensure consistency of the operating points and the satisfaction of external constraints like minimum guaranteed throughputs for transmitted data or maximum power consumptions. Typically, those requirements are ensured on average [13] or for the worst case (robust formulations) [23]. The first approach may perform poorly due to the significant variability in the realizations of the uncertain parameters, while the latter may be too conservative for the same reasons.

Recent models of risk provide sound mathematical foundation and theoretical advantages to decision makers. Models of risk aim at capturing the impact of low-probability but high-cost events. Several approaches to model risk are known: expected utility functions which impose a non-linear transformation of the realizations of the random quantity in question [24]; rank-dependent (dual) utility functions (also called distortions) which transform the probabilities of events [25], [26]; chance constraints or constraints on probabilities of events such as Value-at-Risk constraints [27]–[29], coherent measures of risk [30]–[33], and stochastic-order constraints [34]–[36]. The latter two approaches are subject to intensive research recently because they allow to put an emphasis on the tails of the distributions or to shape the distribution in an efficient way. Risk optimization is widely used in finance and insurance. In wireless networking problems, risk-averse optimization may provide more reliable transmissions, while considering rare events, such as traffic bursts or loss of signal.

In this paper, we are concerned with the problem of deter-

W.-J. Ma, C. Oh, Y. Liu, and M. M. Zavlanos are with the Dept. of Mechanical Engineering and Materials Science, Duke University, Durham, NC 27708, USA {wann.jiun.ma, chanwook.oh, yang.liu.261, michael.zavlanos}@duke.edu. D. Dentcheva is with the Dept. of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ 07030, USA darinka.dentcheva@stevens.edu. This work is supported by ONR under grant #N000141410479.

mining optimal routing decisions in wireless networks as well as selecting optimal locations for the sink nodes, also called access points, that ensure reliable delivery of information from a set of source nodes to the selected infrastructure of access points. We formulate the problem as a risk-averse network flow problem with binary variables corresponding to the status of the sinks, namely, selected or not. By relaxing the integer constraints, we reformulate the problem as a linear optimization problem, which we solve in a distributed way using the ADAL (Accelerated Distributed Augmented Lagrangian) method [37] that was recently developed by the authors to solve optimization problems with convex separable objectives and linear coupling constraints. The contribution of this work lies in formulating a risk-averse model for this problem using coherent measures of risk and in proposing a distributed numerical method for its solution. An additional contribution are the experimental results that demonstrate the ability of the proposed method to effectively deal with large variations in the quality of the wireless channel, that mean-value and worst-case formulations can not handle.

II. Problem Definition

A. Optimal Wireless Networking

In this paper, we employ the communication model proposed in [12], [13]. Specifically, consider a wireless network consisting of J source nodes indexed by $i \in \mathcal{J} = \{1, \dots, J\}$ that route information to K Access Points (APs) indexed by $i \in \mathcal{K} = \{J+1, \dots, J+K\}$. We assume that APs only receive and do not transmit any data. Point-to-point connectivity is modeled through a rate function R_{ij} that determines the amount of information that is transmitted from node $i \in \mathcal{J}$ and is correctly decoded by node $j \in \mathcal{J} \cup \mathcal{K}$. We assume that direct communication with the APs is not always possible, so the source nodes need to route data to the APs in a multi-hop fashion. Routing of packets is due to routing decisions T_{ij} that represent the fraction of time that node i selects node j as its intended destination. Note that, since the routing variables T_{ij} represent time slot shares they need to satisfy $0 \leq T_{ij} \leq 1$ and $\sum_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij} = 1$. The products $R_{ij}T_{ij}$ denote the point-to-point rate of information that is transmitted by node i and correctly received by node j. Between their generation or arrival from another node and their transmission, packets are stored in a queue. The total rate at which packets arrive at the queue of node i is $\sum_{j \in \mathcal{J}} T_{ji}R_{ji}$. Similarly, the total rate at which packets leave the queue of node i is $\sum_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij}R_{ij}$. Collecting all routing decisions in a matrix $\mathbf{T} \in [0,1]^{J \times (J+K)}$, we can express the end-to-end information rate $r_i(\mathbf{T})$ at node i as

$$r_i(\mathbf{T}) = \sum\nolimits_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij}R_{ij} - \sum\nolimits_{j \in \mathcal{J}} T_{ji}R_{ji}. \tag{1}$$

This is the rate at which node i can add data in the network. Integrity of the communication network requires that the end-to-end rates $r_i(\mathbf{T})$ exceed minimum thresholds $r_{i,\min}$ that capture the rate of data that is directly generated at nodes, i.e., $r_i(\mathbf{T}) \geq r_{i,\min}$ for all $i \in \mathcal{J}$. Let $\mathbf{r}(\mathbf{T}) \in \mathbb{R}^J$ be the vector stacking all rates $r_i(\mathbf{T})$. Introducing concave optimality criteria $f(\mathbf{r}(\mathbf{T}))$ measuring the utility associated with the end-to-end rates $\mathbf{r}(\mathbf{T})$, optimal operating points can then be selected as

$$\begin{aligned} \max_{\mathbf{T} \in [0,1]^{J \times (J+K)}} \quad & f(\mathbf{r}(\mathbf{T})) \\ \text{s.t} \quad & r_{i,\min} \leq r_i(\mathbf{T}) = \sum\nolimits_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij}R_{ij} - \sum\nolimits_{j \in \mathcal{J}} T_{ji}R_{ji}, \\ & \sum\nolimits_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij} = 1, \end{aligned} \tag{2}$$

where the constraints hold for all $i \in \mathcal{J}$. Given the communication rates R_{ij}, problem (2) is convex in the decision variables T_{ij} and can be efficiently solved using available optimization techniques. Examples of utilities in (2) are linear $f(\mathbf{r}(\mathbf{T})) = \sum_{i \in \mathcal{J}} w_i r_i(\mathbf{T})$, logarithmic $f(\mathbf{r}(\mathbf{T})) = \sum_{i \in \mathcal{J}} \log(r_i(\mathbf{T}))$, or min-rate $f(r(\mathbf{T})) = \min_{i \in \mathcal{J}}\{r_i(\mathbf{T})\}$. Linear utilities yield larger rates and favor nodes close to the APs. Concave utility functions yield fairer operating points because they penalize small rates. Problem (2) can also incorporate constraints on the overall traffic in the network to leave room for critical traffic or to ensure a fixed power consumption per unit time. These constraints take the form $\sum_{i \in \mathcal{J}} r_i(\mathbf{T}) \leq r_{\text{total}}$ or $r_i(\mathbf{T}) \leq r_{i,\max}$.

B. The Access Point Selection Problem

Let $\mathcal{K} = \{J+1, \dots, J+K\}$ be a set of possible AP locations distributed in the area where the wireless network is deployed. The optimal AP selection problem consists of determining the minimum number of APs in $\mathcal{K}$, and their corresponding locations, that are necessary to solve problem (2). To do so, introduce binary variables $z_i \in \{0,1\}$ that indicate whether an AP at location $i \in \mathcal{K}$ is selected or not. We use $z_i = 1$ to denote a selected AP and $z_i = 0$ to denote an AP that is not selected. The relation to the routing decisions is ensured by the constraints $0 \leq T_{ij} \leq z_j$ for all $i \in \mathcal{J},\ j \in \mathcal{K}$. Collecting all binary variables in $\mathbf{z} \in \{0,1\}^K$, the feasible set $\mathcal{X}$ takes on the form

$$\begin{aligned} \mathcal{X} = \{(\mathbf{z}, \mathbf{T}) \in \{0,1\}^K \times [0,1]^{J \times (J+K)} \ \mid \ & \sum_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij} = 1, \\ & 0 \leq T_{\ell k} \leq z_k \quad \forall\, i, \ell \in \mathcal{J},\ k \in \mathcal{K}\}. \end{aligned}$$

The AP selection problem can be formulated as a Mixed Integer Program (MIP) as follows

$$\begin{aligned} \max_{(\mathbf{z},\mathbf{T}) \in \mathcal{X}} \quad & \sigma f(\mathbf{r}(\mathbf{T})) - \kappa \sum\nolimits_{j \in \mathcal{K}} z_j \\ \text{s.t} \quad & r_{i,\min} \leq r_i(\mathbf{T}) = \sum\nolimits_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij}R_{ij} - \sum\nolimits_{j \in \mathcal{J}} T_{ji}R_{ji}, \end{aligned} \tag{3}$$

where the constraints hold for all $i \in \mathcal{J}$. Here $\sigma > 0$ and $\kappa > 0$ are parameters allowing for a trade-off between the two parts of the objective: the performance measure and the number of APs.

If the communication rates R_{ij} are known, problem (3) can be solved efficiently using methods of combinatorial optimization, such as branch and bound or branch and cut, where tailored bounding techniques and valid inequalities may be developed alongside suitable decomposition techniques. However, in practice, the rates R_{ij} depend on the signal-to-noise ratio and, thus, they are subject to uncertainty due to multipath fading effects that are difficult to predict [38].

Adopting this point of view, part of the data in problem (3) is represented by random variables, which necessitates the formulation of a proper deterministic equivalent of the stochastic problem (3). We propose to address this challenge using risk-averse optimization which we discuss next.

III. Risk-Averse Access Point Selection

A. Risk-Averse Optimization

Measures of risk are widely used in finance and insurance since the mean–variance model was proposed by Markowitz [39]. In engineering and statistics, measures such as Fano factor [40] or the index of dispersion [41] are of similar spirit. An axiomatic theory of measures of risk is presented in [30], [31], [42]. In more general settings, risk measures are analyzed in [32], [43], [44]; see also [33] and the references therein.

We consider the cost function $L(x,y) : \mathcal{X} \times \mathcal{Y} \to \mathbb{R}$ of the decision variable $x \in \mathcal{X} \subset \mathbb{R}^n$ and a parameter vector $y \in \mathcal{Y} \subset \mathbb{R}^m$. Here the set $\mathcal{X}$ denotes the set of feasible decisions and the parameter is modeled by a random vector Y with realization in the set $\mathcal{Y}$. In our problem setting, we have a preference to smaller outcomes for the random cost $L(x,Y)$. Naturally, the expected value of the cost and its variance are the most frequently used distributional characteristics of $L(x,Y)$. Denote by $F_L(\alpha)$ the right-continuous cumulative distribution function of $L(x,Y)$, namely, $F_L(\alpha) = \mathsf{Pr}[L(x,Y) \le \alpha]$. The Value at Risk, $\mathsf{V@R}_\beta(L)$, is defined as

$$\mathsf{V@R}_\beta(L) := \min\{\alpha \in \mathbb{R} \mid F_L(\alpha) \ge \beta\}. \tag{4}$$

If F_L is continuous, then $\mathsf{V@R}_\beta(L)$ is a real number such that $\mathsf{Pr}[L(x,Y) \ge \mathsf{V@R}_\beta(L)] = 1-\beta$. In general, it is the value of the left-continuous inverse of the distribution function at β, which we denote by $F_L^{(-1)}(\beta)$. The risk here is defined as the probability of the cost $L(x,Y)$ obtaining a large value. Although $\mathsf{V@R}$ and the variance of a random variable are intuitively appealing measures, neither of them reflects our preferences adequately. A *coherent measure of risk* is a monotonic, convex, positively homogeneous functional $\varrho : \mathcal{L}_p(\Omega) \to \overline{\mathbb{R}}$, satisfying the property $\varrho(L+a) = \varrho(L) + a$ for all random variables L and all constants $a \in \mathbb{R}$, where $\mathcal{L}_p(\Omega)$ is the space of random variables with finite p-th moments, $p \in [1,\infty]$, and $(\Omega, \mathcal{F}, P)$ is a probability space (cf. [31]–[33]).

In this paper, we focus on the Average Value at Risk $\mathsf{AV@R}_\beta(L)$ (also called Conditional Value at Risk) as our risk model due to its fundamental role among the coherent measures of risk. It is defined as (c.f., [33, (6.23)]).

$$\mathsf{AV@R}_\beta(L) := \inf_\alpha \Big\{\alpha + \frac{1}{1-\beta}\mathsf{E}[\max(0, L(x,Y)-\alpha)]\Big\}. \tag{5}$$

In our numerical experiments, we also consider the upper mean-semideviation, which is a popular coherent measure of risk, defined as follows:

$$\varrho(L) = \mathsf{E}[L(x,Y)] + \theta\Big(\mathsf{E}\big[\big(L(x,Y) - \mathsf{E}[L(x,Y)]_+^p\big]\Big)^{\frac{1}{p}}, \tag{6}$$

where $p \in [1,\infty)$ and $\theta \in [0,1]$ are modeling parameters.

While definition (5) of $\mathsf{AV@R}$ facilitates optimization, the nature and the name of that risk measure are revealed when relating it to the Lorenz function of the random variable involved. Consider the integrated survival function of the random variable $L(x,Y)$ defined as $\bar{F}_L^{(2)}(\eta) = \int_\eta^\infty \big[1 - F_L(t)\big]\, dt = \mathsf{E}[(L(x,y)-\eta)_+]$. The *upper Lorenz function* $\bar{F}_L^{(-2)} : \mathbb{R} \to \overline{\mathbb{R}}$ introduced in [36] as a counterpart of the absolute Lorenz function (cf. [45]) is defined as

$$\bar{F}_L^{(-2)}(\beta) = \int_\beta^1 \mathsf{V@R}_t(L)\, dt \quad \text{for } 0 < \beta < 1. \tag{7}$$

Additionally, $\bar{F}_L^{(-2)}(1) = 0$, $\bar{F}_L^{(-2)}(0) = \mathsf{E}(X)$, and $\bar{F}_L^{(-2)}(\beta) = -\infty$ for $\beta \notin [0,1]$. The function $\bar{F}_L^{(2)}(\cdot)$ is concave because its derivative is monotonically non-increasing. Recall that, for a convex function $f : \mathbb{R}^n \to \mathbb{R}$, its Fenchel conjugate function, f^*, is defined as $f^*(w) = \sup_v\{\langle v, w\rangle - f(v)\}$.

Theorem 3.1 ([36]): The Fenchel conjugate function of $\bar{F}_L^{(2)}(\cdot)$ is the function $-\bar{F}_L^{(-2)}(\cdot + 1)$. Furthermore, for any $\beta \in (0,1)$, it holds $\mathsf{E}(Z-\eta)_+ - \eta(\beta-1) = \bar{F}_L^{(-2)}(\beta)$, whenever η is a β-quantile of X.

Theorem 3.1 is a counterpart of the conjugate duality relation for the absolute Lorenz curve established in [46, Theorem 3.1]. Now, representation (5) leads to the following:

$$\begin{aligned}(1-\beta)&\mathsf{AV@R}_\beta(L)\\ &= \inf_\alpha \Big\{(1-\beta)\alpha + \mathsf{E}[\max(0, L(x,Y)-\alpha)]\Big\}\\ &= -\sup_\alpha \Big\{(\beta-1)\alpha - \mathsf{E}[\max(0, L(x,Y)-\alpha)]\Big\}\\ &= \bar{F}_L^{(-2)}(\beta - 1 + 1) = \bar{F}_L^{(-2)}(\beta),\end{aligned}$$

where the penultimate equation is due to Theorem 3.1. Therefore, we obtain the representation

$$\mathsf{AV@R}_\beta(L) := \frac{1}{1-\beta}\int_{L(x,y)\ge \mathsf{V@R}_\beta(L)} \mathsf{V@R}_\alpha(L)\, d\alpha. \tag{8}$$

If the distribution function of $L(x,Y)$ is continuous, then we can represent the Average Value at Risk as

$$\mathsf{AV@R}_\beta(L) = \mathsf{E}[L(x,Y) | L(x,Y) \ge \mathsf{V@R}_\beta(L)]. \tag{9}$$

Therefore, $\mathsf{AV@R}_\beta$ represents the conditional expectation of the cost function value $L(x,Y)$ given that it is greater than or equal to $\mathsf{V@R}_\beta(L)$. For a given β, we can minimize $\mathsf{AV@R}_\beta$ by appropriately selecting the decision x. From equations (4) and (9), we see that $\mathsf{AV@R}_\beta$ is larger than $\mathsf{V@R}_\beta$, i.e., $\mathsf{AV@R}_\beta(L) \ge \mathsf{V@R}_\beta(L)$, for all $x \in \mathcal{X}$. In practice, if N realizations $\{y^1, y^2, \cdots, y^N\}$ of the random variable Y are available, then we can use a plug-in estimator in (5) or in (8) to estimate the Average Value at Risk. The estimator is strongly consistent as shown in [47].

To minimize risk while ensuring that the mean value is within some tolerance $\tau \in \mathbb{R}$, we can solve the problem

$$\min_{x\in\mathcal{X}} \quad \mathsf{AV@R}_\beta(L(x,Y)) \quad \text{s.t.} \quad \mathsf{E}[L(x,Y)] \le \tau.$$

Alternatively, we can also solve the problem

$$\min_{x\in\mathcal{X}} \quad \lambda\mathsf{AV@R}_\beta(L(x,Y)) \;+\; (1-\lambda)\mathsf{E}[L(x,Y)],$$

where $0 < \lambda < 1$ is a parameter allowing for a trade-off between the risk and the average loss.

B. Risk-Averse Access Point Selection

Define the rate function

$$G_i(\mathbf{T},\mathbf{R}) = r_{i,\min} + \sum\nolimits_{j\in\mathcal{J}} R_{ji}T_{ji} - \sum\nolimits_{j\in\mathcal{J}\cup\mathcal{K}} R_{ij}T_{ij},$$

where $\mathbf{R} \in [0,1]^{J\times(J+K)}$ is the matrix stacking all the R_{ij} for all $i \in \mathcal{J}$ and $j \in \mathcal{J}\cup\mathcal{K}$. Using (5), the $\mathsf{AV@R}$ of the overflow rate $H_i(\mathbf{T},\mathbf{R}) = (G_i(\mathbf{T},\mathbf{R}))_+$ at level β for source node $i \in \mathcal{J}$ takes on the form

$$\mathsf{AV@R}_\beta\big(H_i(\mathbf{T},\mathbf{R})\big) = \inf_{\alpha_i}\left\{\alpha_i + \frac{1}{1-\beta}\mathsf{E}\big[\big(H_i(\mathbf{T},\mathbf{R})-\alpha_i\big)_+\big]\right\}$$

where $\alpha_i \in \mathbb{R}$ is an auxiliary variable. Using $\mathsf{AV@R}$ we can formulate a risk-averse AP selection problem as

$$\begin{aligned}\min_{\boldsymbol{\alpha},\,(\mathbf{z},\mathbf{T})\in\mathcal{X}} \quad & -\sigma\mathsf{E}[f(r(\mathbf{T}))] + \kappa\sum_{j\in\mathcal{K}} z_j \\ & + \sum_{i\in\mathcal{J}}\Big(\alpha_i + \frac{1}{1-\beta}\mathsf{E}\big[\big(H_i(\mathbf{T},\mathbf{R})-\alpha_i\big)_+\big]\Big) \\ \text{s.t} \quad & \mathsf{E}[G_i(T,R)] \le -\tau, \quad \forall\, i\in\mathcal{J}.\end{aligned} \tag{10}$$

Here $\tau \ge 0$ is a positive constant, $\boldsymbol{\alpha} \in \mathbb{R}^J$ is the vector with components α_i, and the maximization in (3) has now become a minimization. The objective of the optimization problem (10) consists of three parts: network utility, sink selection, and $\mathsf{AV@R}_\beta\big(H_i(T,R)\big)$. By minimizing $\mathsf{AV@R}_\beta\big(H_i(\mathbf{T},\mathbf{R})\big)$, the risk of overflow rates can be minimized.

To solve (10), we rely on samples of $\mathbf{R}$, denoted by $\mathbf{R}^n$ for $n \in \mathcal{N} = \{1,\ldots,N\}$, where N is the sample size. In this case, we can introduce auxiliary variables $\mathbf{v}^n \in \mathbb{R}^J$ to substitute the non-smooth terms $H_i(\mathbf{T},\mathbf{R})$ and add additional inequality constraints $v_i^n \ge G_i(\mathbf{T},\mathbf{R}^n)$ for all $i \in \mathcal{J}$ and $n \in \mathcal{N}$, where $\mathbf{v}^n$ is the vector stacking all v_i^n for all $i \in \mathcal{J}$. Similarly, we introduce variables $\mathbf{u}^n \in \mathbb{R}^J$ and additional inequality constraints to substitute the non-smooth terms $\big(H_i(\mathbf{T},\mathbf{R}^n)-\alpha_i\big)_+$ for all $i \in \mathcal{J}$. This allows us to convert problem (10) to the mixed integer linear program

$$\begin{aligned}\min_{\boldsymbol{\alpha},\,\mathbf{u}\ge 0,\,\mathbf{v}\ge 0,\,(\mathbf{z},\mathbf{T})\in\mathcal{X}} \quad & -\sigma\mathsf{E}[f(\mathbf{r}(\mathbf{T}))] + \kappa\sum\nolimits_{j\in\mathcal{K}} z_j \\ & + \sum_{i\in\mathcal{J}}\left(\alpha_i + \frac{1}{N(1-\beta)}\sum_{n=1}^{N} u_i^n\right) \\ \text{s.t} \quad & \mathsf{E}[G_i(\mathbf{T},\mathbf{R})] \le -\tau, \\ & u_i^n \ge v_i^n - \alpha_i,\;\; v_i^n \ge G_i(\mathbf{T},\mathbf{R}^n),\end{aligned} \tag{11}$$

where the constraints hold for all $i \in \mathcal{J}$ and all $n \in \mathcal{N}$, and $\mathbf{u} \in \mathbb{R}^{JN}$ and $\mathbf{v} \in \mathbb{R}^{JN}$ are the vectors stacking the components $\mathbf{u}^n$ and $\mathbf{v}^n$ for all $n \in \mathcal{N}$, respectively.

To solve problem (11), we can approximate it by a convex optimization problem using sparse optimization methodologies. Specifically, we replace the objective that controls the number of active sinks with a convex sparsity-promoting term (ℓ_1-norm), $\kappa\sum_{j\in\mathcal{K}}\|\mathbf{T}_{AP,j}\|_1$, where $\mathbf{T}_{AP,j} \in [0,1]^{J\times 1}$ is the matrix stacking all incoming flow T_{ij} from source node $i \in \mathcal{J}$ to sink $j \in \mathcal{K}$, and set the binary variables $z_j = 1$ for all $j \in \mathcal{K}$. Then problem (11) becomes

$$\begin{aligned}\min_{\boldsymbol{\alpha},\,\mathbf{u}\ge 0,\,\mathbf{v}\ge 0,\,\mathbf{T}\in\tilde{\mathcal{X}}} \quad & -\sigma\mathsf{E}[f(\mathbf{r}(\mathbf{T}))] + \kappa\sum_{j\in\mathcal{K}}\|\mathbf{T}_{AP,j}\|_1 \\ & + \sum_{i\in\mathcal{J}}\left(\alpha_i + \frac{1}{N(1-\beta)}\sum_{n=1}^{N} u_i^n\right) \\ \text{s.t} \quad & \mathsf{E}[G_i(\mathbf{T},\mathbf{R})] \le -\tau, \\ & u_i^n \ge v_i^n - \alpha_i,\;\; v_i^n \ge G_i(\mathbf{T},\mathbf{R}^n),\end{aligned} \tag{12}$$

where the constraints hold for all $i \in \mathcal{J}$ and all $n \in \mathcal{N}$, and the feasible set for $\mathbf{T}$ is defined as

$$\begin{aligned}\tilde{\mathcal{X}} = \{\mathbf{T} \in [0,1]^{J\times(J+K)} \;\mid\; & \sum_{j\in\mathcal{J}\cup\mathcal{K}} T_{ij} = 1, \\ & 0 \le T_{\ell k} \le 1 \quad \forall\, i,\ell \in \mathcal{J},\; k \in \mathcal{K}\}.\end{aligned}$$

In this new formulation (12), a sink is inactive if all incoming flows to that sink are zero, i.e., if $T_{ij} = 0$ for all source nodes i connected to the sink j. Similarly, a sink is active if there exists a source node i so that $T_{ij} > 0$ for sink j. To minimize the number of active sinks we seek solutions to (12) that are sparse, i.e., have large numbers of zero entries. To do so, we can use a re-weighting iterative procedure with the ℓ_1-norm regularization objective [48].

IV. Distributed Implementation

In this section, we propose a distributed solution to problem (12) based on the ADAL algorithm [37]. ADAL is an iterative algorithm within the augmented Lagrangian framework for solving optimization problems with convex separable objectives and linear coupling constraints in a distributed way. In the context of our network problem, a generic iteration of the method includes the following tasks. First, the nodes solve convex optimization subproblems and transmit their solution to the neighboring nodes. Then, every node updates its primal variables based on the previous iterates and the new solutions to the subproblems. Finally, all nodes update their dual variables and transmit them to their neighboring nodes. Assuming that source node $i \in \mathcal{J}$ controls the routing decisions T_{ij} for all $j \in \mathcal{J}\cup\mathcal{K}$ and all other decision variables in (12) that are local to node i, and that sink node $j \in \mathcal{K}$ controls the routing decisions T_{ij} for all $i \in \mathcal{J}$, there are three types of coupling constraints in (12): the routing sum constraints $\sum_{j\in\mathcal{J}\cup\mathcal{K}} T_{ij} = 1$, for all $i \in \mathcal{J}$, the mean routing flow constraints $\mathsf{E}[G_i(\mathbf{T},\mathbf{R})] \le -\tau$ for all $i \in \mathcal{J}$, and the routing constraints $v_i^n \ge G_i(\mathbf{T},\mathbf{R}^n)$ for all $i \in \mathcal{J}$ corresponding to each sample $n \in \mathcal{N}$. Since ADAL requires equality constraints, we first introduce slack variables $h_i \ge 0$ for all $i \in \mathcal{J}$ and $w_i^n \ge 0$, for all $i \in \mathcal{J},\ n \in \mathcal{N}$ so that the constraints $\mathsf{E}[G_i(\mathbf{T},\mathbf{R})] \le -\tau$ and $v_i^n \ge G_i(\mathbf{T},\mathbf{R}^n)$ can be written as $\mathsf{E}[G_i(\mathbf{T},\mathbf{R})] + \tau + h_i = 0$ and $v_i^n = G_i(\mathbf{T},\mathbf{R}^n) + w_i^n$, respectively. Then, we define Lagrange multipliers $\mu_i \in \mathbb{R}$ associated with the constraints $\sum_{j\in\mathcal{J}\cup\mathcal{K}} T_{ij} = 1$ for all $i \in \mathcal{J}$, Lagrange multipliers $\psi_i \in \mathbb{R}$ associated with the constraints $\mathsf{E}[G_i(\mathbf{T},\mathbf{R})] + \tau + h_i = 0$ for all $i \in \mathcal{J}$, and Lagrange multipliers $\delta_i^n \in \mathbb{R}$ associated with the constraints $v_i^n = G_i(\mathbf{T},\mathbf{R}^n) + w_i^n$ for all $i \in \mathcal{J},\ n \in \mathcal{N}$.

To simplify notation, we can define appropriate matrices $\mathbf{E}_i$ for all $i \in \mathcal{J} \cup \mathcal{K}$ and rewrite the coupling constraints $\sum_{j\in\mathcal{J}\cup\mathcal{K}} T_{ij} = 1$ for all $i \in \mathcal{J}$ as $\sum_{i\in\mathcal{J}\cup\mathcal{K}} \mathbf{E}_i\mathbf{x}_i = \mathbf{1}$, where $\mathbf{x}_i$ is the vector stacking all decision variables of node $i \in \mathcal{J}\cup\mathcal{K}$, and $\mathbf{1}$ is a vector with all ones. Similarly, the coupling constraints $\mathsf{E}[G_i(\mathbf{T},\mathbf{R})] + \tau + h_i = 0$ for all $i \in \mathcal{J}$ and $v_i^n = G_i(\mathbf{T},\mathbf{R}^n) + w_i^n$ for all $i \in \mathcal{J},\ n \in \mathcal{N}$ can be rewritten as $\sum_{i\in\mathcal{J}\cup\mathcal{K}} \mathbf{H}_i\mathbf{x}_i = \mathbf{0}$ and $\sum_{i\in\mathcal{J}\cup\mathcal{K}} \mathbf{F}_i\mathbf{x}_i = \mathbf{0}$, respectively, with appropriate matrices $\mathbf{F}_i$ and $\mathbf{H}_i$ for all $i \in \mathcal{J}\cup\mathcal{K}$, where $\mathbf{0}$ is a vector with all zeros. Finally, we denote by $\boldsymbol{\mu}$, $\boldsymbol{\psi}$, and $\boldsymbol{\delta}^n,\ n \in \mathcal{N}$, the vectors stacking all multipliers μ_i, ψ_i, and $\delta_i^n,\ n \in \mathcal{N}$, for all $i \in \mathcal{J}$, respectively, and denote by $\boldsymbol{\delta}_i$ the vector stacking all δ_i^n for all $n \in \mathcal{N}$.

We consider source and sink nodes separately. For every source node $i \in \mathcal{J}$, the corresponding local augmented Lagrangian $\Lambda_{S,i}$ of ADAL at iteration k is defined as

$$\Lambda_{S,i}(\mathbf{x}_i,\mathbf{x}^k,\boldsymbol{\mu}^k,\boldsymbol{\delta}^k,\boldsymbol{\psi}^k) = \left(\alpha_i + \frac{1}{N(1-\beta)}\sum_{n=1}^{N} u_i^n\right) \quad (13)$$
$$- \sigma\mathsf{E}[f_i(\mathbf{r}(\mathbf{T}))] + (\boldsymbol{\mu}^k)^\top\mathbf{E}_i\mathbf{x}_i + (\boldsymbol{\delta}^k)^\top\mathbf{F}_i\mathbf{x}_i + (\boldsymbol{\psi}^k)^\top\mathbf{H}_i\mathbf{x}_i$$
$$+ \frac{\rho}{2}\|\mathbf{E}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{E}_j\mathbf{x}_j^k - \mathbf{1}\|^2$$
$$+ \frac{\rho}{2}\|\mathbf{F}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{F}_j\mathbf{x}_j^k\|^2 + \frac{\rho}{2}\|\mathbf{H}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{H}_j\mathbf{x}_j^k\|^2,$$

where $\mathbf{x}_j^k$ denotes the vector of decision variables communicated to node i from node j at iteration k and $\mathbf{x}^k, \boldsymbol{\mu}^k, \boldsymbol{\delta}^k, \boldsymbol{\psi}^k$ are the vectors stacking the corresponding decision variables. In (13), we also assume that the utility $\mathsf{E}[f_i(\mathbf{r}(\mathbf{T}))]$ is a separable function, e.g., of linear or logarithmic form, as discussed in Section II-A. The penalty coefficient $\rho \geq 0$ is a parameter which controls the tradeoff between optimization of the cost function and satisfaction of the constraints [37]. Note that, source node i needs access to a subset of the entries of $\mathbf{x}^k, \boldsymbol{\mu}^k, \boldsymbol{\delta}^k, \boldsymbol{\psi}^k$ to formulate the local augmented Lagrangian (13). Specifically, define the graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where an edge (i,j) belongs to $\mathcal{G}$ if and only if $R_{ij} > 0$ for all $i,\ j \in \mathcal{V}$. Source node i needs access to the routing decisions T_{ij}^k from the neighboring sink nodes $j \in \mathcal{K}$ and all the routing decisions that are included in the routing flow constraints of its 1-hop source node in $\mathcal{G}$. Also, source node i needs access to the dual variable μ_i^k corresponding to its sum flow constraint and the dual variables $(\boldsymbol{\delta}_i^k, \psi_i^k)$ corresponding to the routing flow constraints of its 1-hop source node in $\mathcal{G}$.

Similarly, for every sink node $i \in \mathcal{J}$, the local augmented Lagrangian $\Lambda_{AP,i}$ of ADAL at iteration k is defined as

$$\Lambda_{AP,i}(\mathbf{x}_i,\mathbf{x}^k,\boldsymbol{\mu}^k,\boldsymbol{\delta}^k,\boldsymbol{\psi}^k) = \kappa\|\mathbf{T}_{AP,i}\|_1 + (\boldsymbol{\mu}^k)^\top\mathbf{E}_i\mathbf{x}_i \quad (14)$$
$$+ (\boldsymbol{\delta}^k)^\top\mathbf{F}_i\mathbf{x}_i + (\boldsymbol{\psi}^k)^\top\mathbf{H}_i\mathbf{x}_i + \frac{\rho}{2}\|\mathbf{E}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{E}_j\mathbf{x}_j^k - \mathbf{1}\|^2$$
$$+ \frac{\rho}{2}\|\mathbf{F}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{F}_j\mathbf{x}_j^k\|^2 + \frac{\rho}{2}\|\mathbf{H}_i\mathbf{x}_i + \sum_{j\in\mathcal{J}\cup\mathcal{K}}^{j\neq i}\mathbf{H}_j\mathbf{x}_j^k\|^2.$$

Because the entries of $\mathbf{T}_{AP,i}$ are non-negative, the ℓ_1-norm $\kappa\|\mathbf{T}_{AP,i}\|_1$ is equal to the sum of its entries $\kappa\sum_{j\in\mathcal{J}} T_{ji}$. Moreover, note that every sink node needs access to the entries of the primal variables corresponding to the routing constraints of only those source nodes that can route directly to the sink. Also, every sink node only needs access to the entries of the dual variables $(\boldsymbol{\mu}^k, \boldsymbol{\delta}^k, \boldsymbol{\psi}^k)$ corresponding to the constraints of all source nodes that can route directly to the sink.

The group of nodes iteratively solve their assigned subproblems. In particular, for the source node $i \in \mathcal{J}$, the local subproblems are

$$\min_{\mathbf{x}_i\in\mathcal{C}_{S,i}} \Lambda_{S,i}(\mathbf{x}_i,\mathbf{x}^k,\boldsymbol{\mu}^k,\boldsymbol{\delta}^k,\boldsymbol{\psi}^k), \quad (15)$$

where the feasible set is

$$\mathcal{C}_{S,i} = \{(\mathbf{T}_{S,i},\alpha_i,\mathbf{u}_i,\mathbf{v}_i) \in [0,1]^{1\times J}\times\mathbb{R}\times\mathbb{R}^N\times\mathbb{R}^N \mid$$
$$0 \leq T_{ij} \leq 1,\ u_i^n \geq v_i^n - \alpha_i, \quad \forall\, i \in \mathcal{J},\ n \in \mathcal{N}\},$$

where $\mathbf{T}_{S,i} \in [0,1]^{1\times J}$ is the matrix stacking all outgoing flows T_{ij} from source node $i \in \mathcal{J}$ to source node $j \in \mathcal{J}$, and $\mathbf{u}_i$ and $\mathbf{v}_i$ are vectors stacking all u_i^n and v_i^n for all $n \in \mathcal{N}$, respectively. For sink node $i \in \mathcal{K}$, the local subproblems are

$$\min_{\mathbf{x}_i\in\mathcal{C}_{AP,i}} \Lambda_{AP,i}(\mathbf{x}_i,\mathbf{x}^k,\boldsymbol{\mu}^k,\boldsymbol{\delta}^k,\boldsymbol{\psi}^k), \quad (16)$$

where the feasible set is

$$\mathcal{C}_{S,i} = \{\mathbf{T}_{AP,i} \in [0,1]^{J\times 1} \mid 0 \leq T_{ji} \leq 1, \quad \forall\, i \in \mathcal{J}\}.$$

Using the solutions $\hat{\mathbf{x}}_i^k$ to subproblems (15) and (16), each node $i \in \mathcal{J}\cup\mathcal{K}$ updates the primal variable as

$$\mathbf{x}_i^{k+1} = \mathbf{x}_i^k + \eta(\hat{\mathbf{x}}_i^k - \mathbf{x}_i^k), \quad (17)$$

where $\eta > 0$ is a parameter that indirectly controls convergence of ADAL and is required to satisfy $\eta \leq 1/d$, where d is the maximum node degree of the graph $\mathcal{G}$ [37].

Finally, after the primal variable update, each source node $i \in \mathcal{J}$ updates the dual variables as

$$\mu_i^{k+1} = \mu_i^k + \eta\rho\Big(\sum_{j\in\mathcal{J}} T_{ij}^{k+1} + \sum_{j\in\mathcal{K}} T_{ij}^{k+1} - 1\Big), \quad (18)$$

$$\delta_i^{n,k+1} = \delta_i^{n,k} + \eta\rho\Big(-v_i^{n,k+1} + G_i(\mathbf{T}^{k+1},\mathbf{R}^n) + w_i^{n,k+1}\Big), \quad (19)$$

$$\psi_i^{k+1} = \psi_i^k + \eta\rho\Big(\mathsf{E}[G_i(\mathbf{T}^{k+1},\mathbf{R}^n)] + \tau + h_i^{k+1}\Big), \quad (20)$$

for all $n \in \mathcal{N}$. Since each source node controls its own dual variables, the dual variable updates are distributed. Note that there are no dual updates for sink nodes based on the decomposition of the decision variable considered here. As mentioned in Section II-B, re-weighting can be used to further promote sparsity of the solutions obtained form the ℓ_1 regularization method considered here. Due to space limitations, the details are omitted but the approach is similar to [48].

V. Numerical Simulations

In the first simulation, we compare the solution of the centralized risk-averse AP selection problem (11) to a model with individual chance constraints and to a mean-value formulation,

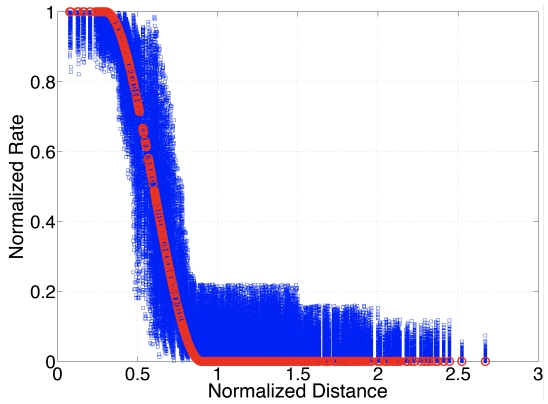


Fig. 1. Empirical distribution of R_{ij}. Red circles represent the nominal values of R_{ij} based on a fixed node configuration. Blue squares represent the samples drawn from a given probability distribution.

both of which are typically used in the literature to handle uncertainty. In particular, we consider the problem

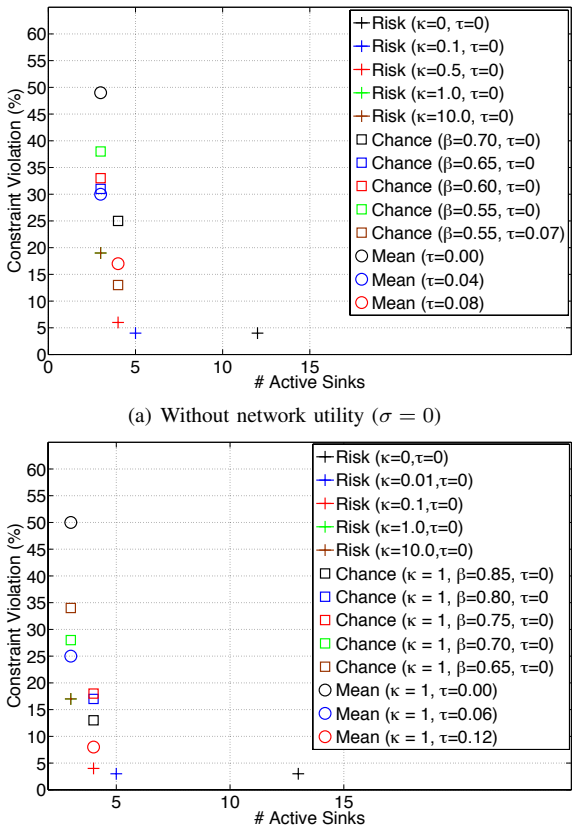
$$\min_{(\mathbf{z},\mathbf{T})\in\mathcal{X}} \quad -\sigma\mathsf{E}[f(r(\mathbf{T}))] + \kappa\sum_{j\in\mathcal{K}} z_j \tag{21}$$
$$\text{s.t.} \quad \mathsf{Pr}\{G_i(\mathbf{T},\mathbf{R}) > -\tau\} \leq (1-\beta) \quad \forall\, i\in\mathcal{J},$$

with confidence level $0 \leq \beta \leq 1$. This formulation was used in [23] to design robust controllers that jointly control mobility and wireless communications in networks of robots. The advantage of using chance constraints is that they can be formulated as second-order cone constraints for normal random variables and the problem can be solved using available convex optimization methods. This assumption limits the application of the approach and leads to very conservative solutions. Additionally, we consider the mean-value formulation

$$\min_{(\mathbf{z},\mathbf{T})\in\mathcal{X}} \quad -\sigma\mathsf{E}[f(r(\mathbf{T}))] + \kappa\sum_{j\in\mathcal{J}} z_j \tag{22}$$
$$\text{s.t.} \quad \mathrm{E}\{G_i(\mathbf{T},\mathbf{R})\} \leq -\tau \quad \forall\, i\in\mathcal{J},$$

where $\tau \geq 0$ is a tolerance. Note that if the communication rates R_{ij} are subject to large variations, then the probability of the event $G_i(\mathbf{T},\mathbf{R}) \geq 0$ may be substantial, despite $\mathsf{E}\{G_i(\mathbf{T},\mathbf{R})\} \leq 0$. Therefore, mean-value approaches are usually ineffective in producing realistic solutions.

Consider a network with 25 source nodes ($J = 25$) and 15 candidate sink nodes ($K = 15$). The reliabilities R_{ij} for all $i \in \mathcal{J}$ and $j \in \mathcal{J}\cup\mathcal{K}$ are generated based on a fixed node configuration. Reliability samples R_{ij}^n are drawn from truncated normal distributions. We consider 100 samples, i.e., $N = 100$ for each R_{ij}. The observations of the samples are shown in Fig. 1. Note that the reliability samples have large variations. The required rates $r_{i,\min}$ are set to 0.45 for all $i \in \mathcal{J}$. Moreover, linear utilities $f(\mathbf{r}(\mathbf{T})) = \sum_{i\in\mathcal{J}} r_i(\mathbf{T})$ are considered in all problems (11), (21), and (22). The solver MOSEK with CVX [49] in Matlab is used to solve problems (11), (21), and (22), in a centralized way. The parameter β in AV@R$_\beta$ is set to 0.9. We check the constraint violations based on the optimal transmission rates T^* and the samples



(a) Without network utility ($\sigma = 0$)

(b) With network utility ($\sigma = 1$)

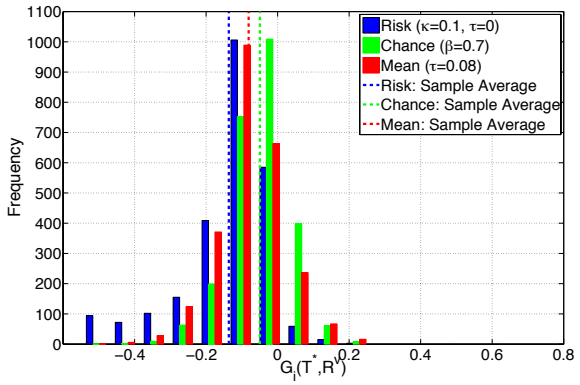
Fig. 2. Plot of the constraint violation versus the number of active sinks with and without considering the network utility. "Risk", "Chance", and "Mean" refer to the formulations (11), (21), and (22), respectively.

R_{ij}^n (shown in Fig. 1). The constraint violation conditions

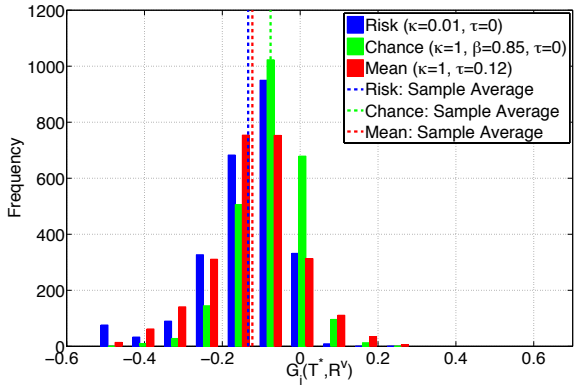
$$\sum_{j\in\mathcal{J}} R_{ij}^n T_{ij}^* + \sum_{j\in\mathcal{K}} R_{ij}^n T_{ij}^* - \sum_{j\in\mathcal{J}} R_{ji}^n T_{ji}^* - r_{i,\min} < 0, \tag{23}$$

are checked one-by-one for all $i \in \mathcal{J}$ and $n \in \mathcal{N}$. The constraint violation ratio is defined as the number of instances in which the condition (23) is true divided by the total number of instances $J \times N$. A sink node $j \in \mathcal{K}$ is activated if $\frac{1}{N}\sum_{n\in\mathcal{N}}\sum_{i\in\mathcal{J}} R_{ij}^n T_{ij}^* \geq 10^{-1}$. A link is established if $\frac{1}{N}\sum_{n\in\mathcal{N}} R_{ij}^n T_{ij}^* \geq 5\times 10^{-2}$. We use the same notation T^* to denote the solutions to problems (11), (21), and (22).

Figs. 2(a) and 2(b) show the results of the AV@R$_\beta$, mean, and chance constraint formulations for $\sigma = 0$ and $\sigma = 1$, respectively. For $\sigma = 0$, problem (22) is infeasible for $\tau > 0.08$ (with 0.01 increment). For the chance constraint formulation, problem (21) is infeasible for $\beta > 0.70$ (with 0.05 increment). For $\sigma = 1$, problem (22) is infeasible for $\tau > 0.12$ (with 0.01 increment). For the chance constraint formulation, problem (21) is infeasible for $\beta > 0.85$ (with 0.05 increment). Also, if we require the sample constraints $G_i(\mathbf{T},\mathbf{R}^n) \leq 0$ to hold for all $i \in \mathcal{J}$ and $n \in \mathcal{N}$ (the worst-case formulation), the problem obtained by replacing the expectation in (22) with the sample constraints is infeasible as well. Figs. 2(a) and 2(b) show that the risk-averse formulation provides a tradeoff between the number of active sinks and the constraint violations. As the weight κ decreases, the constraint violation ratio decreases



(a) Without network utility ($\sigma = 0$)



(b) With network utility ($\sigma = 1$)

Fig. 3. Empirical distributions and sample averages of $G_i(T^*, R^n)$, $i \in \mathcal{J}$, $n \in \mathcal{N}$ for (11), (21), and (22) with and without considering the network utility. "Risk", "Chance", and "Mean" refer to the formulations (11), (21), and (22), respectively.

with a moderate increase in the number of active sinks. By tuning the weight κ, one can increase reliability by increasing the number of active sinks. On the other hand, the mean and chance constraint formulations fail to provide such a tradeoff between the constraint violation and the number of active sinks. These formulations select a very small number of active sinks subject to very high constraint violation ratio. By controlling the parameter β, the risk-averse formulation can provide a wide spectrum of risk preferences, from the most robust but conservative ($\beta = 1$) to risk-neutral ($\beta = 0$).

Figs. 3(a) and 3(b) show the empirical distribution of $G_i(\mathbf{T}^*, \mathbf{R}^n)$ for $i \in \mathcal{J}$ and $n \in \mathcal{N}$ for problems (11), (21), and (22). Since $J = 25$ and $N = 100$, there are total 2500 instances for each formulation in the empirical distribution plot. In Fig. 3(a), we select $\sigma = 0$, $\kappa = 0.1$, and $\tau = 0$ in (11), $\sigma = 0$ and $\kappa = 1$ in (21) and (22), and $\tau = 0.08$ in (22). The confidence level β and τ are set to 0.7 and 0, respectively, in (21). In Fig. 3(b), we select $\sigma = 1$, $\kappa = 0.01$, and $\tau = 0$ in (11), $\sigma = 1$ and $\kappa = 1$ in (21) and (22), and $\tau = 0.12$ in (22). The confidence level β and τ are set to 0.85 and 0, respectively, in (21). Figs. 3(a) and 3(b) show that by minimizing AV@R, the area under the right tail of the constraint violations $G_i(\mathbf{T}^*, \mathbf{R}^n)$ is reduced (shaped). As the number of positive realizations of $G_i(\mathbf{T}^*, \mathbf{R}^n)$ decreases, the constraint violation ratio decreases.

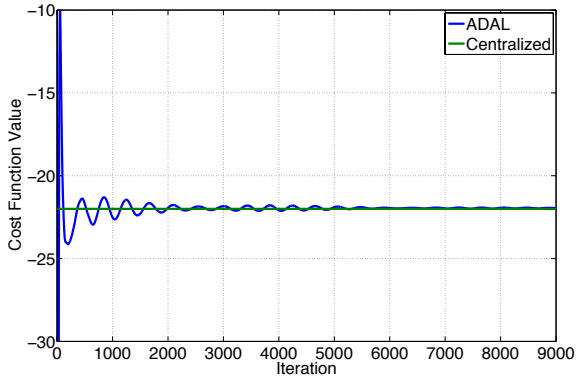
In the second simulation, we verify the proposed distributed method to solve the risk-averse AP selection problem (12). We consider a network with 20 source nodes ($J = 20$) and 15 candidate sink nodes ($K = 15$). As before, the reliabilities R_{ij} for all $i \in \mathcal{J}$ and $j \in \mathcal{N}$ are generated based on a fixed node configuration. Reliability samples R_{ij}^n are drawn from truncated normal distributions. We consider 50 samples, i.e., $N = 50$, for each R_{ij}. The required rates $r_{i,\min}$ are set to 0.15 for all $i \in \mathcal{J}$, and the utility is assumed to be linear as before, i.e., $f(\mathbf{r}(\mathbf{T})) = \sum_{i \in \mathcal{J}} r_i(\mathbf{T})$. The solver MOSEK with CVX [49] in Matlab is used to solve the centralized problem (11) and CPLEX in Matlab is used to solve the primal update in ADAL. The parameters σ, κ, and τ are set to 1, 0.4, and 0, respectively, and β in AV@R is set to 0.9. In ADAL, we set the parameter $\rho = 0.05$. We also apply re-weighting with parameter $\epsilon = 0.1$, as in [48], to enhance sparsity of the solution. Figs. 4(a) and 4(b) verify the convergence of ADAL. Fig. 4(a) shows that the cost function value converges to the one obtained by the centralized method. Fig. 4(b) shows convergence of the constraint violations for the routing sum constraints $\sum_{j \in \mathcal{J} \cup \mathcal{K}} T_{ij} = 1$, for all $i \in \mathcal{J}$, the mean routing flow constraints $\mathsf{E}[G_i(\mathbf{T}, \mathbf{R})] \leq -\tau$ for all $i \in \mathcal{J}$, and the routing constraints derived from each sample $v_{n,i} \geq G_i(\mathbf{T}, \mathbf{R}^n)$, for all $i \in \mathcal{J}$ and $n \in \mathcal{N}$. Figs. 5(a) and 5(b) show the sinks selected and final routing decisions obtained by ADAL and the centralized method, respectively. ADAL selects only one more sink.

VI. Experimental Validation

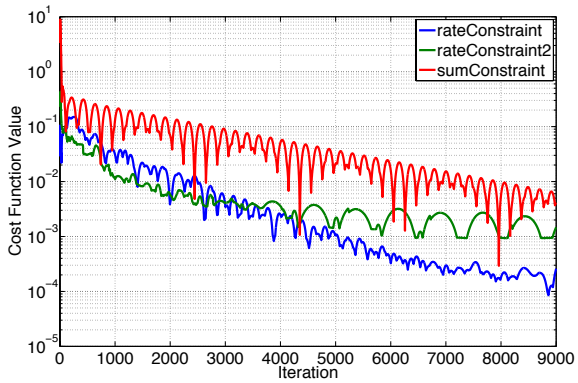
To demonstrate the methods and theoretical findings in this paper, we also conducted real-world experiments using our wireless sensor platform developed in our lab. This platform consists of a sensor, a radio, and a computing device. Detailed descriptions of the hardware and software components can be found in Appendix B. In our experiments, we assume that sufficient performance can be achieved by enforcing the minimum end-to-end rates $r_{i,\min}$ and, therefore, we ignore the expected utility of the routing decisions ($\sigma = 0$). Moreover, since our experimental platform only contains a few wireless nodes, deactivating some of them does not produce meaningful network topologies. Therefore, we do not consider sink selection in our experiments ($\kappa = 0$). Instead, we focus on the proposed risk-averse framework (11) and show that it produces more reliable routing decisions compared to approaches that do not consider risk, such a mean-value formulations (22).

The distributed ADAL algorithm discussed in Section IV was implemented on the wireless nodes using Python to solve problem (11). Problem (22) can be decomposed and solved using ADAL, in a similar way. Implementation details are discussed in Appendix C. For comparison, we also consider the mean-semideviation measure of order one in (6). Following a process similar to that in Section III-B, we can obtain a linear risk-averse network flow problem using this risk measure as:

$$\begin{aligned} \min_{\boldsymbol{\alpha},\, \mathbf{u} \geq 0,\, \mathbf{T} \in \tilde{\mathcal{X}}} \quad & \sum_{i \in \mathcal{J}} \left(\mathsf{E}[G_i(\mathbf{T}, \mathbf{R})] + \frac{1}{N} \sum_{n=1}^{N} u_i^n \right) \\ \text{s.t} \quad & \mathsf{E}[G_i(\mathbf{T}, \mathbf{R})] \leq -\tau, \\ & u_i^n \geq G_i(\mathbf{T}, \mathbf{R}) - \mathsf{E}[G_i(\mathbf{T}, \mathbf{R})], \end{aligned} \tag{24}$$

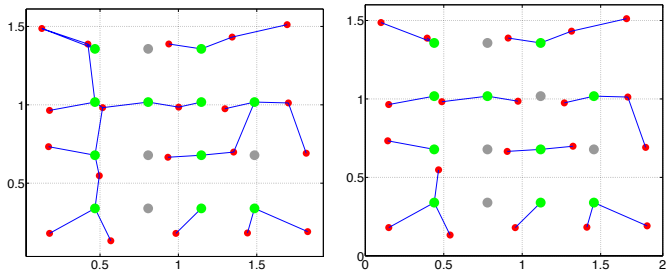


(a) Convergence of cost function.



(b) Convergence of constraint violations.

Fig. 4. Convergence of the distributed ADAL algorithm discussed in Section IV. Fig. 4(a): Cost function versus the iteration of ADAL; Fig. 4(b): Log of maximum constraint violation for the routing sum constraint ("sumConstraint"), mean routing flow constraints ("rateConstraint"), and the sample flow constraints ("rateConstrainr2").



(a) ADAL (b) Centralized

Fig. 5. Network configuration and selected sinks. Red nodes are the source nodes. Green and grey nodes are the active sinks and inactive sinks, respectively.

where the constraints hold for all $i \in \mathcal{J}$ and all $n \in \mathcal{N}$ and, as before, we omit the network utility and sink selection objectives. Problem (24) can be decomposed and solved using ADAL, in a similar way as discussed in Section IV.

A. Experimental Characterization of Channel Uncertainty

Initial experiments were run to verify the operation of the radio sensors as they respond to different environments and situations (e.g., hallways and rooms, people walking past). Changes in RSSI values were sensed by the radio hardware and sent to the connected computer, where they were logged.

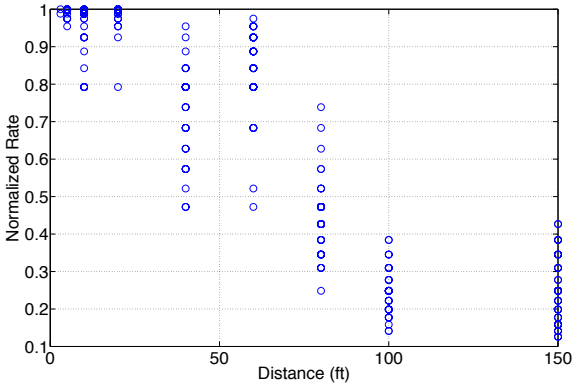


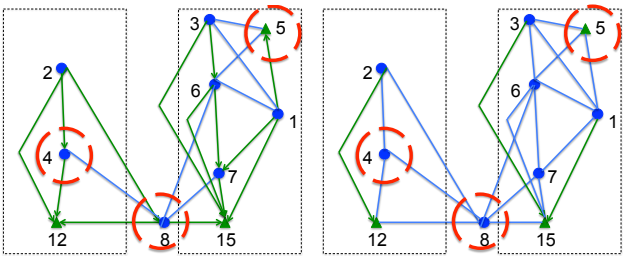
Fig. 6. Samples of the rate function R_{ij} collected in an indoor environment. There are $N = 100$ samples at each distance (many samples are overlapping).

Fig. 7. Wireless sensor network configuration (first floor, Hudson Hall, Duke University). There are 3 sink and 7 source nodes. Blue circle: source node; green triangle: sink node.

RSSI measurements serve as samples to be used in the risk optimization methods proposed in this paper. Specifically, we performed an experiment to measure the RSSI values in an indoor environment. We took $N = 100$ RSSI samples for each channel based on a fixed node configuration. Sensors simultaneously transmitted messages and each message received by a sensor contained RSSI values. We used these RSSI values to calculate the normalized rates R_{ij} using equations (26) and (27) in Appendix A. The R_{ij} samples are shown in Fig. 6. Note that the rate samples have large variations and their distribution resembles Fig. 1.

B. Validation of Risk-Averse Routing Decisions

We carried out an experiment in which nodes relay data from the source nodes to sink nodes based on a unicast transmission protocol that employs those optimal routing decisions obtained from (11), (22), and (24). There are 7 source nodes and 3 sink nodes in the workspace. The sensor configuration is shown in Fig. 7. We allow nodes 1, 2, 3 to generate packets so that $r_{i,\min} = 0.2$ for nodes 1, 3, and $r_{i,\min} = 0.25$ for node 2. All other source nodes help to relay packets to sink nodes. We select three nodes (sink 5, source 8, and source 4) and introduce sporadic interference on the incoming and outgoing links for these three nodes. Interference events are rare, but significantly increase the variance of RSSI samples. Figs. 8(a) and 8(b) depict the final routing decisions obtained by the risk-neural (22) and risk-averse (11) method, respectively. Formulation (24) produces routing decisions that are similar



(a) Risk-neutral routing (b) Risk-averse routing

Fig. 8. Routing decisions for the node configuration shown in Fig. 7. There are 3 sink and 7 source nodes. Blue circles and green triangles correspond to the sources and sinks, respectively. Green arrows represent the routing decisions. Blue lines represent links that have not been selected. Red circles indicate nodes where interference was introduced.

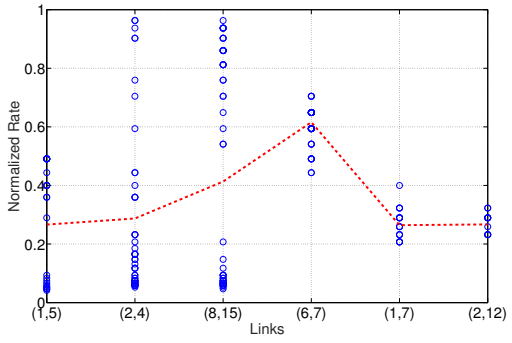


Fig. 9. Samples of the rate function R_{ij} for the links of interest. Red dotted line indicates the average rate values for the corresponding links. There are $N = 100$ samples for each link (many samples are overlapping).

to (11) and are not shown. Notice that the risk-neutral method returns solutions that utilize the unreliable links (1,5) and (2,4) that are subject to interference. This is because, even though the frequency of low RSSI values for these links is high, their average RSSI values are also high, as shown in Fig. 9. Since links (1,5) and (2,4) are unreliable, we expect a higher number of failed transmissions so that the queue sizes of the nodes will be larger compared to the risk-averse formulations and the number of packets received at the sinks will be lower. This effect is indeed observed in Figs. 10(a), 10(b), and 10(c) that show the queue size of each node over a 10 minute experiment. Fig. 11 shows that within 10 minutes of experiment, the routing decisions obtained by the risk-averse methods allow the sinks to receive more packets than the ones obtained by the risk-neutral method; specifically, 721, 729, and 699 packets for the AV@R, semi-deviation, and risk-neutral methods, respectively. Fig. 12 shows that the risk-averse methods reduce the number of positive realizations of $G_i(\mathbf{T}^*, \mathbf{R}^n)$ compared to the risk-neutral method; compare also to Figs. 3 in Section V.

VII. Conclusions

We proposed a method to select the optimal set of APs and routing decisions in wireless communication networks while at the same time, taking into account uncertainty in the wireless channel due to multipath fading effects. Incorporating a coherent measure of risk into the AP selection problem leads to solutions which are more reliable compared to the one obtained by the mean-value approach and less conservative than the worst-case approach. The proposed risk-averse algorithm can be implemented in a distributed way. We carried out real world experiments to demonstrate the effectiveness of the proposed method to deal with large variations in the quality of wireless channels.

References

[1] E. M. Royer and C.-K. Toh, "A review of current routing protocols for ad-hoc mobile wireless networks," *IEEE Personal Communications*, vol. 6, no. 2, pp. 46–55, April 1999.

[2] E. Montijano, J. I. Montijano, and C. Sagues, "Adaptive consensus and algebraic connectivity estimation in sensor networks with chebyshev polynomials," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, Dec 2011, pp. 4296–4301.

[3] M. Franceschelli, A. Gasparri, A. Giua, and C. Seatzu, "Decentralized estimation of laplacian eigenvalues in multi-agent systems," *Automatica*, vol. 49, no. 4, pp. 1031 – 1036, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0005109813000307

[4] L. Sabattini, N. Chopra, and C. Secchi, "Decentralized connectivity maintenance for cooperative control of mobile robotic systems," *The International Journal of Robotics Research*, vol. 32, no. 12, pp. 1411–1423, 2013.

[5] M. M. Zavlanos and G. J. Pappas, "Potential fields for maintaining connectivity of mobile networks," *IEEE Transactions on Robotics*, vol. 23, no. 4, pp. 812–816, August 2007.

[6] ——, "Distributed connectivity control of mobile networks," *IEEE Transactions on Robotics*, vol. 24, no. 6, pp. 1416–1428, December 2008.

[7] M. Ji and M. Egerstedt, "Distributed coordination control of multiagent systems while preserving connectedness," *IEEE Transactions on Robotics*, vol. 23, no. 4, pp. 693–703, Aug 2007.

[8] M. M. Zavlanos, M. B. Egerstedt, and G. J. Pappas, "Graph theoretic connectivity control of mobile robot networks," *Proceedings of the IEEE: Special Issue on Swarming in Natural and Engineered Systems*, vol. 99, no. 9, pp. 1525–1540, September 2011.

[9] A. Ephremides, "Energy concerns in wireless networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 4, pp. 48–59, August 2002.

[10] R. Zhang and M. Alouini, "Channel-aware inter-cluster routing protocol for wireless ad-hoc networks," in *Proc. of the 6th International Symposium on Communication Theory and Applications*, Ambleside, UK, July 2001, pp. 46–51.

[11] D. DeCouto, D. Aguayo, J. Bicket, and R. Morris, "A high-throughput path metric for multihop wireless routing," in *Proc. ofInternational ACM Conference on Mobile Computing and Networking*, San Diego, CA, September 2006, pp. 134–146.

[12] A. Ribeiro, Z.-Q. Luo, N. D. Sidiropoulos, and G. B. Giannakis, "Modelling and optimization of stochastic routing for wireless multihop networks," in *Proc. 26th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM)*, Anchorage, Alaska, May 2007, pp. 1748–1756.

[13] M. M. Zavlanos, A. Ribeiro, and G. J. Pappas, "Network integrity in mobile robotic networks," *IEEE Transactions on Automatic Control*, vol. 58, no. 1, pp. 3–18, January 2013.

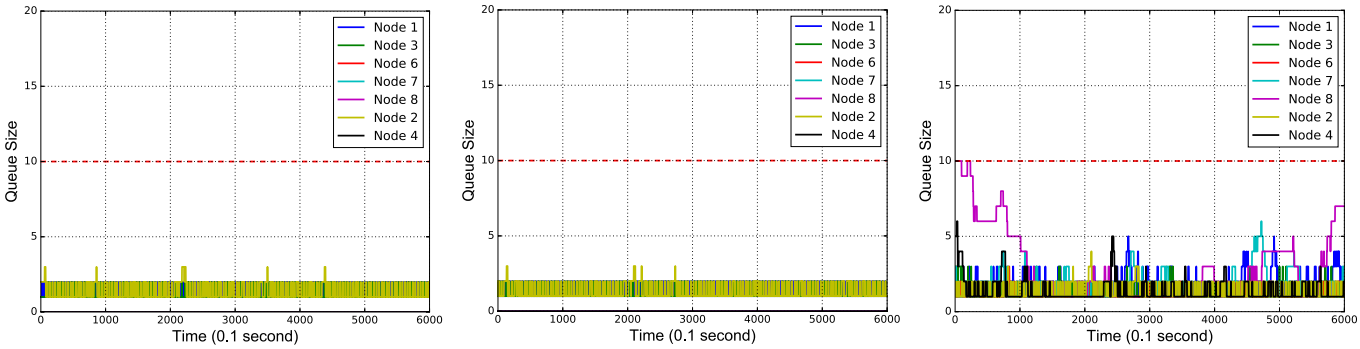
[14] X. Lin, N. B. Shroff, and R. Srikant, "A tutorial on cross-layer optimization in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1452–1463, August 2006.

[15] A. Eryilmaz and R. Srikant, "Joint congestion control, routing, and mac for stability and fairness in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1514–1524, August 2006.

[16] X. Wang and K. Kar, "Cross-layer rate optimization for proportional fairness in multihop wireless networks with random access," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 8, pp. 1548–1559, August 2006.

[17] Y. Yi and S. Shakkottai, "Hop-by-hop congestion control over a wireless multi-hop network," *IEEE/ACM Transactions on Networking*, vol. 15, no. 1, pp. 1548–1559, February 2007.

[18] M. Chiang, S. H. Low, R. A. Calderbank, and J. C. Doyle, "Layering as optimization decomposition: A mathematical theory of network architectures," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 255–312, January 2007.



(a) Average-Value-at-Risk (b) Mean-Semideviation (c) Risk-Neutral

Fig. 10. Real-time queue sizes of all source nodes $i \in \mathcal{J}$ based on the routing decisions obtained by the solution of the risk-averse problems (11) and (24) for $\sigma = \kappa = 0$, and the risk-neutral problem (22).

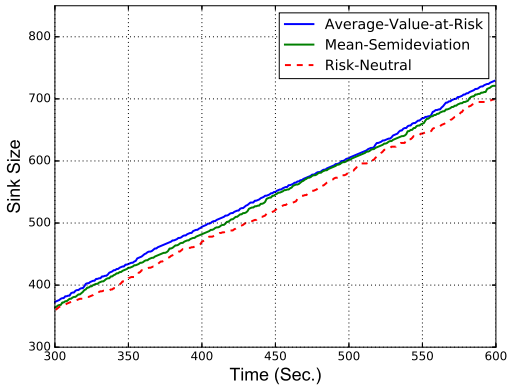


Fig. 11. Real-time sink sizes of all the sink nodes $i \in \mathcal{J}$ for both the risk-averse and risk-neutral methods.

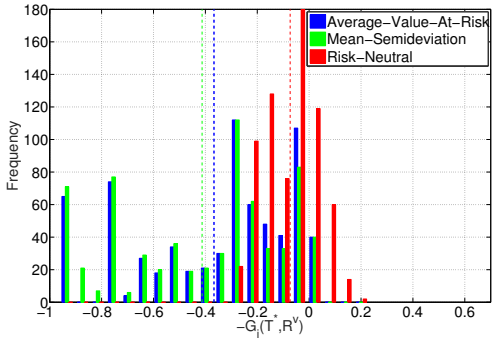


Fig. 12. Empirical distributions and sample averages of $G_i(T^*, R^n)$, $i \in \mathcal{J}$, $n \in \mathcal{N}$ for AV@R, semi-deviation, and risk-neutral formulations.

[19] M. J. Neely, E. Modiano, and C. E. Rohrs, "Dynamic power allocation and routing for time-varying wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 1, pp. 89–103, January 2005.

[20] H. Lundgren, E. Nordstrom, and C. Tschudin, "The gray zone problem in ieee 802.11b based ad hoc networks," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 6, no. 3, pp. 104–105, July 2002.

[21] Y. Mostofi, A. Gonzales-Ruiz, A. Ghaffarkhah, and D. Li, "Characterization and modeling of wireless channels for networked robotic and control systems - a comprehensive overview," in *Proceedings of 2009 International Conference on Intelligent Robots and Systems*, St Louis, MO, October 2009.

[22] J. Fink, N. Michael, A. Kushleyev, and V. Kumar, "Experimental characterization of radio signal propagation in indoor environments with application to estimation and control," in *Proceedings of the International Conference of Intelligent Robot Systems*, St Louis, MO, October 2009.

[23] J. Fink, A. Ribeiro, and V. Kumar, "Robust control for mobility and wireless communication in cyber physical systems with application to robot teams," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 164–178, Jan 2012.

[24] J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944.

[25] J. Quiggin, "A theory of anticipated utility," *Journal of Economic Behavior and Organization*, vol. 3, pp. 225–243, 1982.

[26] ——, *Generalized Expected Utility Theory - The Rank-Dependent Expected Utility Model*. Dordrecht: Kluwer, 1993.

[27] A. Prékopa, *Stochastic Programming*. Dordrecht, Boston: Kluwer Acad. Publ., 1995.

[28] D. Dentcheva, A. Prékopa, and A. Ruszczyński, "Concavity and efficient points of discrete distributions in probabilistic programming," *Mathematical Programming*, vol. 89, pp. 55–77, 2000.

[29] D. Dentcheva and G. Martinez, "Regularization methods for optimization problems with probabilistic constraints," *Mathemathical Programming, Ser. A*, vol. 138, pp. 223–251, 2013.

[30] M. Kijima and M. Ohnishi, "Mean-risk analysis of risk aversion and wealth effects on optimal portfolios with multiple investment opportunities," *Ann. Oper. Res.*, vol. 45, pp. 147–163, 1993.

[31] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, pp. 203–228, 1999.

[32] H. Föllmer and A. Schied, "Convex measures of risk and trading constraints," *Finance and Stochastics*, vol. 6, pp. 429–447, 2002.

[33] A. Shapiro, D. Dentcheva, and A. Ruszczynski, *Lectures on stochastic programming: modeling and theory*. MOS-SIAM Series on Optimization, 2014, vol. 16.

[34] D. Dentcheva and A. Ruszczyński, "Optimization with stochastic dominance constraints," *SIAM Journal on Optimization*, vol. 14, pp. 548–566, 2003.

[35] D. Dentcheva and E. Wolfhagen, "Optimization with multivariate stochastic dominance constraints," *SIAM Journal on Optimization*, vol. 25, no. 1, pp. 564–588, 2015.

[36] D. Dentcheva and G. Martinez, "Two-stage stochastic optimization problems with stochastic ordering constraints on the recourse," *European Journal of Operational Research*, vol. 219, no. 1, pp. 1–8, 2012.

[37] N. Chatzipanagiotis, D. Dentcheva, and M. M. Zavlanos, "An augmented lagrangian method for distributed optimization," *Mathematical Programming*, vol. 152, no. 1–2, pp. 405–434, August 2015.

[38] R. Shorey, A. Ananda, M. C. Chan, and W. T. Ooi, *Mobile, Wireless and Sensor Networks: Technology, Applications and Future Directions*. New York: Wiley, 2006.

[39] H. M. Markowitz, *Mean–Variance Analysis in Portfolio Choice and Capital Markets*. Oxford: Blackwell, 1987.

[40] U. Fano, "Ionization yield of radiations. ii. the fluctuations of the number of ions," *Physical Review*, vol. 72, no. 1, p. 26, 1947.

[41] D. Cox and P. Lewis, "The statistical analysis of series of events," 1966.

[42] R. Rockafellar, S. Uryasev, and M. Zabarankin, "Generalized deviations in risk analysis," *Finance and Stochastics*, vol. 10, pp. 51–74, 2006.

[43] A. Ruszczyński and A. Shapiro, "Optimization of convex risk functions," *Mathematics of Operations Research*, vol. 31, pp. 433–452, 2006.
[44] ——, "Conditional risk mappings," *Mathematics of Operations Research*, vol. 31, pp. 544–561, 2006.
[45] O. Lorenz, "Methods of measuring concentration of wealth," *Journal of the American Statistical Association*, vol. 9, pp. 209–219, 1905.
[46] W. Ogryczak and A. Ruszczyński, "Dual stochastic dominance and related mean risk models," *SIAM Journal on Optimization*, vol. 13, pp. 60–78, 2002.
[47] D. Dentcheva, G. J. Stock, and L. Rekeda, "Mean-risk tests of stochastic dominance," *Statistics & Decisions International mathematical journal for stochastic methods and models*, vol. 28, no. 2, pp. 97–118, 2011.
[48] N. Chatzipanagiotis and M. M. Zavlanos, "Distributed scheduling of network connectivity using mobile access point robots," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1333–1346, Dec 2016.
[49] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

APPENDIX A
UNCERTAINTY IN THE WIRELESS CHANNEL

In practice, the rates R_{ij} depend on the signal-to-noise ratio and, thus, they are subject to uncertainty [38]. This is because received signal strength in wireless communications is typically affected by path loss due to the distance between the transmitter and the receiver, shadowing due to the presence of obstacles in the environment, and multipath fading due to reflections and refractions of the electromagnetic waves. While path loss and shadowing effects can be captured using predictive models, multipath fading is difficult to predict.

In this work, we consider simple radios that do not perform rate or power adaptation. Then, the communication rate R_{ij} is a function of the packet error rate of the channel, which in turn is a function of the signal-to-noise ratio [38]. Specifically, let $x_i \in \mathbb{R}^3$ and $x_j \in \mathbb{R}^3$ denote the 3-D positions of nodes i and j, respectively. One model of channel uncertainty is discussed in [23], where the received signal strength (RSSI) is represented as

$$P_R(x_i, x_j) = L_0 - 10n \cdot \log\big(\|x_i - x_j\|\big) - W(x_i, x_j) - F, \quad (25)$$

where L_0 denotes the measured power at a distance d_0 from the source, n is a path loss exponent, $W(x_i, x_j)$ is a non-smooth function that models shadowing effects, and F is a zero-mean Gaussian random variable with variance σ_F^2 that models fading effects. Using the received signal strength we can approximate the packet error rate by

$$p_{\text{error}}\big(P_R(x_i, x_j)\big) = \text{erfc}\left(\sqrt{k \frac{P_R(x_i, x_j)}{P_{N_0}}}\right), \quad (26)$$

where $\text{erfc}(x)$ is the complementary error function, k is a constant associated with the particular coding and modulation scheme, and P_{N_0} is the noise power. Thus, a probabilistic model for the communication rates R_{ij} can be obtained by multiplying the information rate of transmitted packets R_0 with the probability of successful decoding as

$$R_{ij}(x_i, x_j) = R_0\Big(1 - p_{\text{error}}\big(P_R(x_i, x_j)\big)\Big). \quad (27)$$

Obviously, the uncertainty in the multipath fading effects F is propagated to the communication rates R_{ij}, which become themselves random quantities. In fact, the communication rates R_{ij} are known to experience very large variations in practice [23]. Note that the communication model (25)–(27) assumes that the random variable F is normal. In this case, using sophisticated models of risk to control uncertainty is not necessary as it is sufficient to control the expectation and variance. This is consistent with early approaches to optimal wireless networking under uncertainty [13], [22]. Nevertheless, in practice, fading effects are not normally distributed (one reason is that RSSI can not be negative) and so the rates R_{ij} are also not normal. For this reason, here we do not make any normality assumptions and, in fact, we do not rely on equation (25) to obtain RSSI samples. Instead, we employ a data-driven approach where we directly measure the channel and use these channel measurements in (26)–(27) to obtain samples of R_{ij}.

Fig. 13. Wireless sensor platform: Texas Instruments SensorTag, Raspberry Pi with a 7" touchscreen, and accessories.

```
New Measurement
-------------------------
BAR: Pressure=1005.37 hPa
BAR: Temp=33.00 C
OPT: Light=0.88 lux
HDC: Temp=32.64 C
HDC: Humidity=5.38 %RH
TMP: Ambient=32.687 C
TMP: Object=33.968 C
MPU Gyro: X=3.53 deg/sec
MPU Gyro: Y=-5.51 deg/sec
MPU Gyro: Z=-0.87 deg/sec
MPU Acc: X=0.10 G
MPU Acc: Y=0.04 G
MPU Acc: Z=-0.85 G
```

Fig. 14. Output of unicast messages (including environment's pressure, temperature, humidity, etc.) generated at each source node $i \in \mathcal{J}$.

APPENDIX B
WIRELESS NETWORK

A. Wireless Node

For our experiments, we implement communication networks based on the IEEE 802.15.4 protocol which specifies the physical layer and media access control for low-rate wireless personal area networks (LR-WPANs). Currently, we use Texas Instruments SensorTag (TI CC2650 ARM wireless microcontroller) for development and testing in our lab. SensorTag includes an IEEE 802.15.4 radio to broadcast messages to the other nodes in the network and has the ability to create mesh networks. The CC2650 device embedded in a Sensortag is a member of the CC26xx family of cost-effective, ultralow

power, 2.4-GHz RF devices, and has very low active RF and micro-controller current. The CC2650 device contains a 32-bit ARM Cortex-M3 processor that runs at 48 MHz as the main processor with 128KB of In-System Programmable Flash memory and a rich peripheral feature set that includes a unique ultralow power sensor controller. This sensor controller is ideal for interfacing external sensors and for collecting analog and digital data autonomously while the rest of the system is in sleep mode. Thus, the CC2650 device is ideal for applications within a whole range of products including industrial, consumer electronics, and medical. The sensor is relatively small (about the size of a credit card) and contains a long-term coin cell battery with 1 year battery life or can be powered through its micro-USB connection to a computer.

The SensorTag requires Texas Instruments Debug DevPack, which adds debug capability to our SensorTag for software development. The Debug DevPack is smaller than the SensorTag itself and can be integrated with the sensor and connected to a computer via a micro-USB cable.

Two different computer hardware systems are also being developed and tested: a basic desktop running Ubuntu 14.04, and a Raspberry Pi with a 7" touchscreen running Raspbian. Both systems are capable of interfacing through serial port via micro-USB cables with the sensor hardware, as well as running software for communication routing and numerical optimization algorithms.

B. Software

The software used on all sensor devices is Contiki OS. This is an open source operating system which can run on low power hardware such as our SensorTag and provides functions to help with communication, including broadcast and unicast commands. The broadcast command sends broadcast packets to all nearby sensors. The unicast program looks similar to the broadcast program, however, a destination address is needed to transmit packets from a pre-designed origin to a destination node. An example of output of unicast message is shown in Fig. 14. Contiki OS supports many wireless standards including IPv6, 6LoWPAN, and Zigbee, which may be used in testing. Contiki OS applications are written in standard C language and interface with Python code which runs on the computer hardware and is used to take in data pertaining to the quality of the communication and the radios that communicate and determine routing decisions. Decoupling of radio code from computer code allows the mix-and-match of hardware that we use.

The main solver (along with the numeric module SciPy, a Python package) used to execute the optimization algorithm is CVXOPT. This is a free software package for convex optimization based on the Python programming language. It can be used with the interactive Python interpreter, on the command line by executing Python scripts, or integrated in other software via Python extension modules.

APPENDIX C
ALGORITHM IMPLEMENTATION

To implement ADAL in the wireless network, we need information from each node's 2-hop neighbors. To keep track of these packets, we introduce a variable "hop" that has a value equal to two when a packet is generated. Whenever a packet travels through a communication link, the hop number decreases by one and accordingly, the maximum number of links that a packet can travel through is two. For example, a packet generated at node 6 in Fig. 7 will have hop number two when it is first generated. After it travels through the link (6,5) and is received by node 5, the packet's hop number will become one. Then, after it travels through the link (5,1), the hop number will become zero and the packet will no longer be transmitted. This mechanism ensures that a packet does not travel further than two link, which eliminates unnecessary data transmission in the network.

Implementation of the ADAL algorithm also requires the iterations to be synchronized across the nodes in the network. We implement a synchronization mechanism that relies on the following two phases: communication and computation phases. In the communication phase, each node transmits and receives packets from its neighbors. Each packet carries a flag that contains the iteration index of the ADAL algorithm during which the variables in this packet were updated. Every node periodically checks if the flags received from its neighbors contain the current ADAL iteration index. If the flags of all neighbors are current, then the node moves on to the computation phase and performs a primal or dual update, as per the ADAL algorithm. On the other hand, if the flags of all neighbors are are not current, then the node remains in the communication phase. When a node moves on to the computation phase, it also updates its flag, i.e., by increasing its value by one. We require that each node sends packets containing the current and previous flags. By sending the previous flag, we make sure that all the nodes will eventually move on to the next iteration even when some nodes update slower than others.

The packets communicated in the proposed network contain the primal and dual valuables discussed in Section IV. Standard elements of a packet in Contiki OS are declared as different types of integers. There are four types of integers considered here: 16 bits unsigned, 16 bits signed, 32 bits unsigned, and 32 bits signed integers, where each integer variable can account for different ranges of integers. Since a packet in Contiki can only contain integers, each node needs to transform floating numbers used in the algorithm to integers for communication. We perform this floating point number to integer transformation by multiplying each primal or dual variable by a factor and dropping the decimal digits. When a node receives a packet containing primal and dual variables in an integer format, it divides the variables by the same factor and transforms the variables back to floating point numbers. Then, the floating point numbers are used in the algorithm. Depending on the value of the multiplication factor, the ADAL algorithm may suffer from quantization errors during the floating point to integer transformation. The maximum number of bytes allowed in each packet in Contiki OS is 128 Bytes. Therefore, to use 32 bit integers (uint32 and int32) to maintain the accuracy of all the primal and dual variables, every message needs to be split into to four packets.